CLAIMS

We claim:

1. A method of transforming a Web document from a first format into a second format, comprising:

retrieving a copy of the Web document wherein the Web document comprises at least one element that is delimited and identified by at least one tag within the Web document;

parsing the Web document to create a first data structure comprised of a first hierarchical organization of elements from the Web document;

conducting a semantic analysis of the elements in the data structure; and

re-arranging the elements in the first data structure based upon the semantic analysis to form a second data structure comprised of a new hierarchical organization of elements from the Web document, wherein the new hierarchical organization differs from the first hierarchical organization.

2. A method as defined in claim 1, additionally comprising:

receiving information regarding a user device that requested the Web document; and

creating a device-specific version of the Web document using the second data structure, the device-specific version of the Web document comprised of at least some of the elements in the second data structure, wherein the device-specific version of the Web document is tailored for display on the user device that requested the Web document and is organized according to the new hierarchical organization.

3. A method as defined in claim 2, wherein the information regarding the user device includes memory capacity, display screen size, and data transmission bandwidth.

4. A method as defined in claim 2, wherein the device-specific version of the Web document is divided into discrete data fragments and wherein each data fragment is tailored to fit within data bandwidth capabilities of the user device, memory capabilities of the user device, and display capabilities of the user device.

5. A method as defined in claim 4, wherein the device-specific version of the Web document includes a top level data fragment that represents a top level summary of the Web document.

6. A method as defined in claim 2 wherein the device-specific version of the Web document is written in a markup language that can be interpreted by the user device.

7. A method as defined in claim 1, wherein the Web document comprises descriptive markup language code, and wherein parsing the Web document comprises identifying elements in the Web document based upon the location of the tags in the code and creating a node in the hierarchical structure for each element.

8. A method as defined in claim 7, wherein the descriptive markup language comprises the HyperText Markup Language (HTML), the Extensible Markup Language (XML) , or the Extensible Hypertext Markup Language (XHTML).

9. A method as defined in claim 1, wherein re-arranging the first data structure includes deleting at least some of the elements from the hierarchical structure.

10. A method as defined in claim 1, wherein re-arranging the first data structure includes adding new elements to form the second data structure.

11. A method as defined in claim 1, wherein re-arranging the first data structure includes merging a first element and a second element from the hierarchical structure into a single element.

12. A method as defined in claim 1, wherein conducting a semantic analysis of the elements in the data structure includes analyzing each of the elements in the hierarchical data structure, beginning with elements in a lowermost level in the hierarchical data structure and then analyzing the elements in a level above the lowermost level.

13. A method as defined in claim 1, additionally comprising analyzing the structural arrangement of the elements in the first data structure including examining the location of elements in the data structure with respect to other elements in the data structure.

14. A method as defined in claim 1, wherein semantically analyzing the elements in the first data structure includes determining whether any of the elements are headers.

15. A method as defined in claim 1, wherein semantically analyzing the elements in the first data structure includes determining whether any of the elements are list items.

16. A method as defined in claim 1, wherein semantically analyzing the elements in the data structure comprises categorizing each of the data elements into a predefined category based upon a set of rules and appending an identifier to each data element to identify the category of the data element.

17. A method as defined in claim 14, wherein the first data structure is re-arranged according to the category of the data element.

18. A method of converting a Web page from a first format into a second format, comprising:

identifying page elements in the Web page;

creating a native hierarchical arrangement having nodes that each correspond to a Web page element from the Web page;

performing a structural and semantic analysis on the native hierarchical arrangement according to a set of rules, wherein the semantic analysis comprises examining the relative

location and meaning of each element in the native hierarchical arrangement and identifying nodes for deletion from the hierarchical structure; and

creating a transformed hierarchical arrangement based upon the structural and semantic analysis, wherein the transformed hierarchical arrangement takes into account the relative location and meaning of the elements in the native hierarchical arrangement.

19. A method as defined in claim 18, additionally comprising:

creating at least one transformed Web page comprising Web page elements from the transformed hierarchical arrangement, the Web page elements being arranged according to a hierarchy that corresponds to the transformed hierarchical arrangement.

20. A method as defined in claim 19, wherein the at least one transformed Web pages each have a data size that is tailored to fit within a memory capacity, display screen size, and data transmission bandwidth of a user device that requests the Web page.

21. A method as defined in claim 20, wherein at least one of the transformed Web pages includes a table of contents for the transformed Web pages.

22. A method as defined in claim 18, wherein the native Web page format comprises a HyperText Markup Language, Extensible Markup Language (XML), or Extensible Hypertext Markup Language (XHTML) format.

23. A method as defined in claim 18, wherein the predefined Web page elements comprise elements that are identified by HyperText Markup Language tags.

24. A method as defined in claim 18, wherein at least some of the predefined Web page elements comprise links that point to additional Web pages.

25. A method as defined in claim 18, wherein the method further comprises receiving a request for a Web page and providing the transformed Web page in response to the request.

26. A method as defined in claim 18, wherein the native hierarchical arrangement includes plural levels, and wherein semantic analysis is conducted level-by-level for each level in the native hierarchical arrangement.

27. A method as defined in claim 18, wherein the Web page elements are identified using tags in the Web page.

28. A method as defined in claim 18, wherein each node in the hierarchical arrangement is associated with an identifier that corresponds to the tag for the element associated with the node.

29. A method of transforming a Web document, comprising:

retrieving a native format version of the Web document, the Web document including at least one element that is delimited by at least one tag in the Web document, wherein the native format version of the Web document is not suitable for interpretation and display by a user device that requested the Web document;

performing an analysis of the elements of the Web document, the analysis taking into account semantics of the elements and a structural arrangement of the elements;

rearranging the elements as a result of the analysis to generate a hierarchical data structure that represents the Web document;

generating a user device format version of the Web document based upon the hierarchical data structure, wherein the user device format version of the Web document is suitable for interpretation and display by the user device that requested the Web document.

30. A method as defined in claim 29, additionally comprising:

receiving information regarding a user device that requested the Web document, the information including memory capacity, display screen size, and data transmission bandwidth, wherein the user device format version of the Web document is divided into discrete data fragments and wherein each data fragment is tailored to fit within the memory capacity, data transmission bandwidth, and display screen size of the user device.

31. A method as defined in claim 30, wherein the user device format version of the Web document includes a top level data fragment that represents a top level summary of the Web document.

32. A system that transforms a Web document from a first format into a second format, the system comprising one or more processors that execute program instructions and receive a data set, wherein the program instructions are executed to cause the processor to:

retrieve a copy of the Web document wherein the Web document comprises at least one element that is delimited and identified by tags within the Web document;

parse the Web document to create a first data structure comprised of a first hierarchical organization of elements from the Web document;

conducts a semantic analysis of the elements in the data structure; and

re-arrange the elements in the first data structure based upon the semantic analysis to form a second data structure comprised of a new hierarchical organization of elements from the Web document, wherein the new hierarchical organization differs from the first hierarchical organization.

33. A system as defined in claim 31, wherein the program instructions are further executed to cause the processor to:

receive information regarding a user device that requested the Web document, the information including memory capacity, display screen size, and data transmission bandwidth; and

create a device-specific version of the Web document using the second data structure, the device-specific version of the Web document comprised of at least some of the elements in the second data structure, wherein the device-specific version of the Web document is tailored for display on the user device that requested the Web document and is organized according to the new hierarchical organization.

34. A system as defined in claim 33, wherein the device-specific version of the Web document is divided into discrete data fragments and wherein each data fragment is tailored to fit within data bandwidth capabilities of the user device, memory capabilities of the user device, and display capabilities of the user device.

35. A program product for use in a computer system that executes program steps recorded in a computer-readable media to perform a method for transforming a Web document from a first format into a second format, the program product comprising:

a recordable media;

a program of computer-readable instructions executable by the computer system to perform operations comprising:

retrieving a copy of the Web document wherein the Web document comprises at least one element that is delimited and identified by tags within the Web document;

parsing the Web document to create a first data structure comprised of a first hierarchical organization of elements from the Web document;

conducting a semantic analysis of the elements in the data structure; and

re-arranging the elements in the first data structure based upon the semantic analysis to form a second data structure comprised of a new hierarchical organization of elements from the Web document, wherein the new hierarchical organization differs from the first hierarchical organization.

36. A system that converts a Web page from a first format into a second format, the system comprising one or more processors that execute program instructions and receive a data set, wherein the program instructions are executed to cause the processor to:

identify page elements in the Web page;

create a native hierarchical arrangement having nodes that each correspond to a Web page element from the Web page;

perform a structural and semantic analysis on the native hierarchical arrangement according to a set of rules, wherein the semantic analysis comprises examining the relative location and meaning of each element in the native hierarchical arrangement and identifying nodes for deletion from the hierarchical structure; and

create a transformed hierarchical arrangement based upon the structural and semantic analysis, wherein the transformed hierarchical arrangement takes into account the relative location and meaning of the elements in the native hierarchical arrangement.

37. A program product for use in a computer system that executes program steps recorded in a computer-readable media to perform a method for converting a Web page from a first format into a second format, the program product comprising:

a recordable media;

a program of computer-readable instructions executable by the computer system to perform operations comprising:

identifying page elements in the Web page;

creating a native hierarchical arrangement having nodes that each correspond to a Web page element from the Web page;

performing a structural and semantic analysis on the native hierarchical arrangement according to a set of rules, wherein the semantic analysis comprises examining the relative location and meaning of each element in the native hierarchical arrangement and identifying nodes for deletion from the hierarchical structure; and

creating a transformed hierarchical arrangement based upon the structural and semantic analysis, wherein the transformed hierarchical arrangement takes into account the relative location and meaning of the elements in the native hierarchical arrangement.

38. A system that transforms a Web document, the system comprising one or more processors that execute program instructions and receive a data set, wherein the program instructions are executed to cause the processor to:

retrieve a native format version of the Web document, the Web document including at least one element that is delimited by at least one tag in the Web document, wherein the native format version of the Web document is not suitable for interpretation and display by a user device that requested the Web document;

perform an analysis of the elements of the Web document, the analysis taking into account semantics of the elements and a structural arrangement of the elements;

rearrange the elements as a result of the analysis to generate a hierarchical data structure that represents the Web document;

generate a user device format version of the Web document based upon the hierarchical data structure, wherein the user device format version of the Web document is suitable for interpretation and display by the user device that requested the Web document.

39. A program product for use in a computer system that executes program steps recorded in a computer-readable media to perform a method for transforming a Web document, the program product comprising:

a recordable media;

a program of computer-readable instructions executable by the computer system to perform operations comprising:

retrieving a native format version of the Web document, the Web document including at least one element that is delimited by at least one tag in the Web document, wherein the native format version of the Web document is not suitable for interpretation and display by a user device that requested the Web document;

performing an analysis of the elements of the Web document, the analysis taking into account semantics of the elements and a structural arrangement of the elements;

rearranging the elements as a result of the analysis to generate a hierarchical data structure that represents the Web document;

generating a user device format version of the Web document based upon the hierarchical data structure, wherein the user device format version of the Web document is suitable for interpretation and display by the user device that requested the Web document.

40. A system that transforms a Web document from a first format into a second format, comprising:

a parser that parses a Web document that comprises at least one element that is delimited and identified by at least one tag within the Web document to create a first data structure comprised of a first hierarchical organization of elements from the Web document;

a semantic content analyzer that conducts a semantic analysis of the elements in the data structure; and

a transformer that re-arranges the elements in the first data structure based upon the semantic analysis to form a second data structure comprised of a new hierarchical organization of elements from the Web document, wherein the new hierarchical organization differs from the first hierarchical organization.